

A statistical model for *in vivo* neuronal dynamics

Simone Carlo Surace^{1,2,*}, Jean-Pascal Pfister²

1 Department of Physiology, University of Bern, Bern, Switzerland

2 Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zurich, Switzerland

*** E-mail: surace@pyl.unibe.ch**

Abstract

Cortical neurons are constantly active. Even in the absence of an explicit stimulus, cortical neurons are spontaneously active and display large fluctuations of their membrane potentials. The increasing amount of intracellular recordings of spontaneous activity as well as the increasing number of theories which critically rely on a characterization of spontaneous activity calls for a proper quantification of spontaneous intracellular dynamics. Here we propose a statistical model of spontaneous activity which is very flexible and remains tractable. More specifically, we propose a doubly stochastic process where the subthreshold membrane potential follows a Gaussian process and the spike emission intensity depends nonlinearly on the membrane potential as well as the previous spiking history. We first show that the model has a rich dynamical repertoire since it can capture arbitrary subthreshold autocovariances, firing-rate adaptations as well as arbitrary shapes of the action potential. We then show that this model can be efficiently learned without overfitting. We finally show that this model can be used to characterize and therefore precisely compare a wide range of intracellular *in vivo* recordings from various animals and recording conditions.

Author Summary

The intracellular activity of neurons, given in the form of a voltage trace, contains a wealth of information, which can be characterized by ad-hoc statistical measures of interest, such as mean firing rate, interspike interval distribution or autocorrelation function. However, such a characterization is far from complete, and provides insufficient information in order to allow reproducing synthetic samples of activity in simulations. Here, we designed a statistical model which characterizes the intracellular activity of a neuron by capturing all its important statistical features, without making any assumptions on what gave rise to the activity itself. It describes the structure of the subthreshold membrane potential, the shape of the action potential, the coupling between subthreshold activity and the spiking output, and lastly, the adaptive nature of the spiking. The model is a stochastic process, and so one can throw a die and sample activity which looks different every time it is sampled, while its statistical properties are fixed and given by the parameter setting. When those parameters are obtained by fitting the model to a given *in vivo* recording, the model can generate a synthetic activity consistent with the one recorded.

Introduction

Cortical neurons of awake animals display a high degree of spontaneous activity. From the first intracellular recordings performed in awake cats [1, 2] to more recent recording in cats [3], monkey [4] and mice [5] and even in freely behaving rats [6], it has been shown that the membrane potential displays large fluctuations and is very rarely at its resting potential. Some recent finding in the cat visual cortex have also suggested that the statistical properties of spontaneous activity is comparable to the neuronal dynamics when the animal is exposed to natural images [7]. Similar findings have been shown with

extracellular recordings in the ferret [8]. So the increase amount of intracellular data of awake animals as well the need to compare in a rigorous way the data under various recording conditions call for a proper quantification of spontaneous activity.

Another reason why a precise model of spontaneous activity is urgently needed is that there are several theories that have been proposed that critically depend on statistical properties of spontaneous activity. For example Berkes et al. validate their Bayesian approach of the visual system by comparing the spontaneous activity with the averaged evoked activity [8]. Another Bayesian theory on STP proposed the idea that STP acts as a Bayesian estimator of the presynaptic membrane potential [9]. To validate this theory, it is also necessary to quantify the spontaneous activity.

The aim of our study is to develop a stochastic model of spontaneous activity. In order to be powerful, this model has to be flexible enough in order to capture the large diversity of neuronal dynamics while remaining tractable such that learning the parameters is feasible in a reasonable time. Simple models such as integrate-and-fire types of model with white noise are certainly simple and tractable, but their are not flexible enough to capture the typical long-tailed autocorrelations of spontaneous activity [7]. At the other end of the spectrum, Hodgkin-Huxley types of model with colored noise have certainly a richer dynamical repertoire, but training them remains challenging [10, 11]. Here we develop a new stochastic model for spontaneous activity which is flexible and tractable.

More precisely, we propose a doubly stochastic process where the subthreshold membrane potential follows a Gaussian process and the firing intensity is expressed as a non-linear function of the membrane potential. Since we further include refractoriness and adaptation mechanisms, our model, which we call the Adaptive Gaussian Point Emission process (AGAPE), can be seen as an extension of a Cox process [12].

We first validate the maximum likelihood learning procedure by showing that the learned model is consistent with the one that has been used to generate the data. We then show that the same learning procedure can be applied to intracellular *in vivo* data from various different neurons. We also show that the level of complexity in the model is appropriate, i.e. we show that it does not overfit the data. We finally show that the model captures distinct dynamical features of spontaneous and evoked activity and across multiple brain areas and animals. We conclude the paper by discussing the relationship to other models, the extensions of the AGAPE as well as its potential applications.

Results

Here we present a statistical model of the subthreshold membrane potential and firing pattern of a single neuron *in vivo*. See fig. 1A for such an *in vivo* membrane potential recording. We call the model the Adaptive GAussian Point Emission process (or for short, the AGAPE process) since the subthreshold membrane potential follows a Gaussian process and since the spike emission process is adaptive. A more formal definition of the model will be given in the section below. Then we will show a range of different results. 1) The model is flexible and supports arbitrary autocorrelation structures and adaptation kernels. Therefore, the range of possible statistical features is very large. 2) The model is efficiently learnable and the learning procedure is validated on synthetic data. 3) The model can be fit to *in vivo* datasets. 4) The model is not overly complex, meaning the model performs better even on data which was not used for training compared to simpler versions of the model. Finally, 5) the model is a strong tool to study differences in intracellular dynamics of the same cell during different recording conditions, such as when it is spontaneously active or when a stimulus is presented, thereby providing important biological insights.

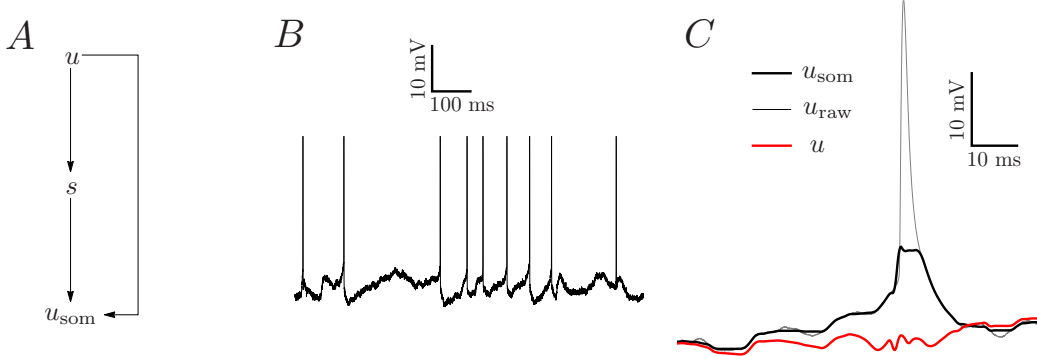


Figure 1. (A) The basic architecture of the model: the subthreshold membrane potential u gives rise to spikes s . The observable u_{som} is a combination of the subthreshold potential and the spikes. (B) A sample *in vivo* membrane potential trace from an intracellular recording of a neuron in HVC of a Zebra Finch. (C) The separation of the recorded membrane potential into sub- and suprathreshold components.

Definition of the AGAPE model

In order to describe the sub- and suprathreshold activity of a neuron, we have to separate the two layers. The subthreshold activity forms an analog signal, whereas the spikes are fairly well approximated by point-like events. The separation of the two quantities calls for lines to be drawn when the action potential starts and where it ends. It is well-known that action potentials have a high degree of stereotypy. We therefore remove from the recorded signal a stereotypic component by subtraction, but this component is not set by hand (e.g. one could remove the spike-triggered average membrane potential, running the risk of removing too much), but instead learned by the model. The resulting membrane potential is then the signal of interest which couples to the spiking mechanism to produce point-like events.

Let us formally introduce the relevant quantities. The AGAPE process defines a probability distribution over the somatic membrane voltage trace $u_{\text{som}}(t)$ and the spike train $s(t) = \sum_{i=1}^{n_s} \delta(t - \hat{t}_i)$ where \hat{t}_i , $i = 1, \dots, n_s$ are the spike times. See Fig. 1B for a graphical representation of the model structure. We assume that the measured membrane voltage as a function of time $u_{\text{som}}(t)$ is given by

$$u_{\text{som}}(t) = u_r + u(t) + u_{\text{spike}}(t), \quad (1)$$

where u_r is a constant (the reference potential), $u(t)$ denotes a stochastic function drawn from a stationary Gaussian process (GP) [13]

$$u \sim \mathcal{GP}[0, k(t - t')] \quad (2)$$

with covariance function $k(t - t')$, and $u_{\text{spike}}(t)$ is the spike shape contribution (see Fig. 1C) which consists of the convolution of a stereotypical spike shape kernel α with the spike train $s(t)$, i.e.

$$u_{\text{spike}}(t) = (\alpha * s)(t) = \int_{-\infty}^{\infty} \alpha(t') s(t - t') dt'. \quad (3)$$

Here, we have made the separation of subthreshold and suprathreshold layers, in that whatever is stereotypic and triggered by the point-like spikes $s(t)$ is attributed to $u_{\text{spike}}(t)$, and the rest belongs to the fluctuating signal $u(t)$. Note that this model could be easily extended by including an additional term in eq. (1) which depends on the input current e.g. a linear filter of the input (see also Discussion). However since this input current was not accessible in our recordings, its contribution was assumed to be part of

$u(t)$. Note also that even though the convolution in eq. (3) extends to the future of t , $u_{\text{spike}}(t)$ only affects the appearance of the total somatic membrane potential, without having an influence on the statistics of spike arrival times. Those are described in the following. We want to introduce the coupling between the subthreshold potential $u(t)$ and the spiking output, as well as adaptive effects associated with spike generation. These effects are summarized by an instantaneous firing rate $r(t)$, which is computed from the value of the membrane potential at time t , $u(t)$, and the previous spike history as

$$r(t) = g[A(t) + \beta u(t)], \quad A(t) = \int_0^\infty \eta(t') s(t - t') dt', \quad (4)$$

where $\beta \geq 0$ is the coupling strength between u and the spikes, and $A(t)$ is the adaptation variable which is the convolution of an adaptation kernel η with the past spike train. The function g is called gain function, and here we use an exponential one, i.e. $g[A(t) + \beta u(t)] = r_0 e^{A(t) + \beta u(t)}$. Other functional forms such as rectified linear or sigmoidal could be used depending on the structure of the data. However, this choice has important implications on the learnability of the model. Using this rate, the probability density for s on an interval $[0, T]$, conditioned on u can be written as

$$p(s|u) = \exp\left(-\int_0^T r(t) dt\right) \prod_{i=1}^{n_s} r(\hat{t}_i). \quad (5)$$

The parameter β connects the subthreshold membrane potential u to the rate fluctuations. The magnitude of the rate fluctuations depend on the variance σ^2 of u , and therefore we sometimes use $\beta\sigma$ as a measure for the effective coupling strength. When $\beta > 0$ the quantity $\theta(t) = -A(t)/\beta$ can be regarded as a soft threshold variable which is modulated after a spike, and $u(t) - \theta(t)$ is the effective membrane potential relevant for the spike generation. This spiking process is a point process which generalizes the log Gaussian Cox process, i.e. when $A = 0$, eq. (5) describes an inhomogeneous Poisson process with rate $g[\beta u(t)]$.

The model has a rich dynamical repertoire

The AGAPE provides a flexible framework which can be adjusted in complexity to model a wide range of data. While for the datasets presented here a covariance function was used which consists of a sum of multiple Ornstein-Uhlenbeck (OU) kernels, the Gaussian Process (GP) allows for arbitrary covariance functions to be used. It allows of course simple exponentially decaying autocorrelation (as produced by a leaky integrate-and-fire neuron), but it can produce also more interesting covariance functions such as power-law covariances which are found in visual cortex [7] or subthreshold oscillations which are found in [14].

The model is also able to reproduce a wide range of firing statistics. A common measure of firing irregularity is the coefficient of variation (C_V , the ratio of standard deviation and mean) of the inter-spike interval distribution. It has been proven that for any Cox process $C_V \geq 1$ [15]. Indeed, in the absence of adaptation, i.e. for $\eta = 0$ the AGAPE model is a Cox process and the inter-arrival times (inter-spike intervals) of this process have a distribution which is influenced by the coupling between the subthreshold potential and the firing rate ($\beta\sigma$ in dimensionless units, where σ is the standard deviation of the subthreshold GP) and the autocorrelation structure of the GP. Therefore, the inter-spike interval C_V is expected to be a function of these parameters which has values greater than one. To illustrate this, we sample synthetic data from a simple version of the AGAPE where the subthreshold potential u is an OU process with time-constant τ . As shown in Fig. 2A, the C_V is an increasing function of the membrane time-constant τ , baseline firing rate r_0 , and dimensionless coupling parameter between membrane potential and firing rate $\beta\sigma$ and the range of the C_V extends from 1 to ≈ 8 within a range of $\beta\sigma \in [0, 2]$ and $r_0\tau \in [2^{-2}, 2^8]$. When adaptation is included in the model, firing statistics change in a significant way, altering also the C_V of the inter-spike interval distribution. For simplicity, we assume an

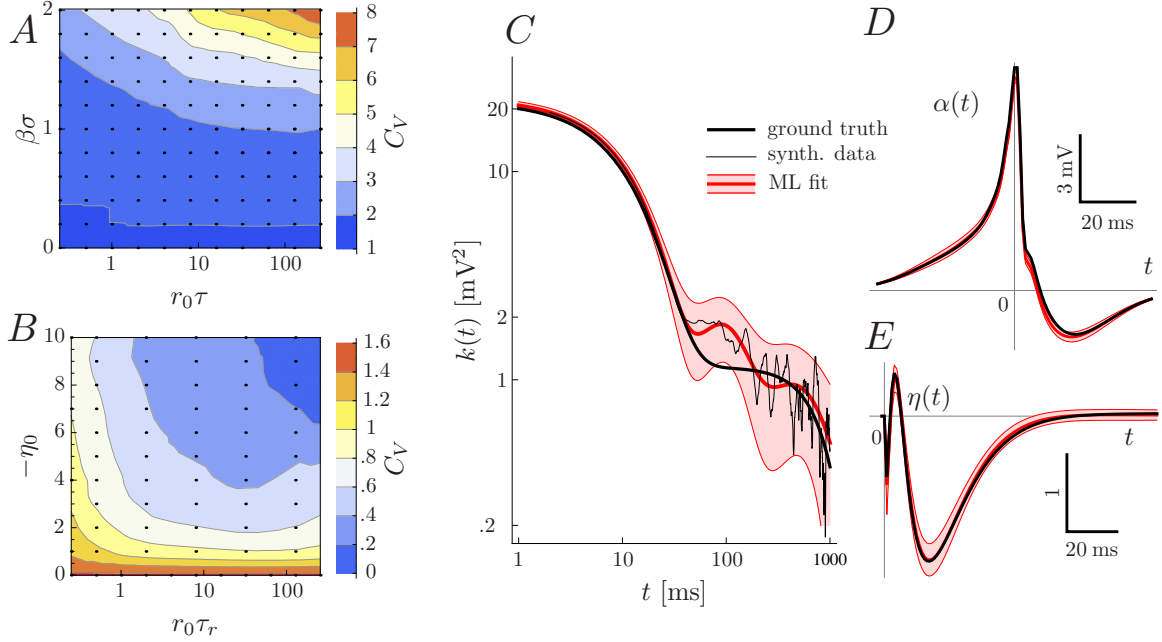


Figure 2. The model has a rich dynamical repertoire (A,B) and can be correctly learned on synthetic data (C-E). (A,B) The coefficient of variation (C_V) of the inter-spike interval distribution is computed for parameter values shown as black dots and then linearly interpolated. (A) The C_V of a simple version of the AGAPE (OU+inhomogeneous Poisson process) as a function of the model parameters (membrane time-constant τ , baseline firing rate r_0 and coupling $\beta\sigma$). (B) C_V of an exponentially adaptive process with fixed membrane time-constant, firing rate and coupling ($\beta\sigma = r_0\tau = 1$) as a function of the parameters describing adaptation, namely adaptation strength η_0 and time-constant τ_r . (C,D,E) Synthetic data is sampled from the model with GP (C), spike-shape (D) and adaptation (E) kernels as depicted in black, and $r_0 = 60$ Hz, $\beta = 0.05$ mV^{-1} . Then the AGAPE is fitted to the synthetic data by maximum likelihood (ML). The ML estimates (red) are seen to lie within two standard deviations (shaded regions, estimated by means of the Cramér-Rao bound) from the ground truth.

exponential adaptive kernel, i.e. $\eta(t) = -\eta_0 e^{-t/\tau_r}$. The same general behavior is seen, with C_V increasing both as a function of $\beta\sigma$ and $r_0\tau$. However, the range of values now also covers the interval $(0, 1)$ which is not accessible by the Cox process but which is observed in many neurons across the brain [16]. In order to see the influence of the parameters of the adaptation mechanism, we fix $\beta\sigma = r_0\tau = 1$ and plot the C_V as a function of $r_0\tau_r$ and η_0 (see Fig. 2B). Within the parameter region shown, the C_V spans values from 0.1 up to 1.6.

The model can be learned efficiently

The coupling of the GP and spiking mechanism through the spike shape and nonlinear coupling parameter β makes maximum likelihood parameter estimation a complex, non-convex problem. Moreover, the evaluation of the GP likelihood of n samples, where $n = \mathcal{O}(10^5)$ comes at a high computational cost. Two important techniques make the parameter learning both tractable and fast: the first is the use of the circulant approximation of the GP covariance matrix which makes evaluation of the likelihood function fast. The second is the use of an alternating learning algorithm which replaces the non-convex optimization in the full parameter space with two convex optimizations and a non-convex one in suitable parameter subspaces. Those two techniques are further described here.

Efficient likelihood computation

The above expressions for the subthreshold membrane potential dynamics (eq. (2)) as well as for the spike emission process (eq. (5)) define the joint log-likelihood $\log p(u_{\text{som}}, s)$, which has two terms, namely

$$\log p(u_{\text{som}}, s) = \log p_s(u_{\text{som}}) + \log p(s|u_{\text{som}}). \quad (6)$$

The log-likelihood function is evaluated in its discrete-time form with n time points separated by a time-step Δt . The GP variable u (which leads to u_{som} through eq. (3)) is multivariate Gaussian distributed with a covariance matrix $K_{ij} = k(t_i - t_j)$, where $t_i = i\Delta t$. The matrix K is symmetric and, by virtue of stationarity, Toeplitz. Evaluation of the GP likelihood requires inversion of K , which is computationally expensive (the time required to invert a matrix typically scales with n^3). For this reason we approximate this Toeplitz matrix by the circulant matrix which minimizes the Kullback-Leibler divergence

$$C = \underset{D \text{ circulant}}{\operatorname{argmin}} \mathcal{D}_{\text{KL}} [\mathcal{N}(m, K) || \mathcal{N}(m, D)] \quad (7)$$

between the two multivariate Gaussian distributions with the same mean but different covariance matrices. This optimization problem can be solved by calculating the derivative of $\mathcal{D}_{\text{KL}} [\mathcal{N}(m, K) || \mathcal{N}(m, D)]$ with respect to D and using the diagonalization of D by a Fourier transform matrix [17]. After a bit of algebra (see Materials and Methods), denoting $k_i = K_{1i}$ and $k_{n+1} \equiv 0$, the optimal circulant matrix can be written as $C_{ij} = c_{(i-j \bmod n)+1}$, where $i, j = 1, \dots, n$ and

$$c_i = \frac{1}{n} [(n-i+1)k_i + (i-1)k_{n-i+2}]. \quad (8)$$

The replacement of K by C is equivalent to having periodic boundary conditions, which has a small effect under the assumption that the time interval spanned by the data is much longer than the largest temporal autocorrelation length of k . So the first term on the r.h.s. of eq. (6) is a multivariate Gaussian density $\mathcal{N}(0, C)$. The determinant of the covariance matrix C is the product of eigenvalues, which for a circulant matrix are conveniently given by the entries of \hat{c} , the discrete Fourier transform of c . Also the scalar product $u^T C^{-1} u$ can be written in terms of \hat{c} [17]. Together, the first term on the r.h.s. of eq. (6) takes the simple form

$$\log p_s(u_{\text{som}}) = -\frac{1}{2} \sum_{i=1}^n \left(\log(2\pi\hat{c}_i) + \frac{1}{n} \frac{|\hat{u}_i|^2}{\hat{c}_i} \right). \quad (9)$$

The Gaussian component of the membrane potential u is implicitly given by $u = u_{\text{som}} - u_r - \alpha * s$, and \hat{u} is its discrete Fourier transform. The time required to compute $\log p_s(u_{\text{som}})$ is determined by the complexity of the Fourier transform, which is of the order of $n \log n$. This dramatic reduction in complexity (compared to n^3) allows a fast evaluation of the log-likelihood.

The spiking distribution $p(s|u_{\text{som}})$ is approximated by a Bernoulli distribution in each time bin (i.e. $s_i = 1$ if there is a spike at time \hat{t}_j with $t_i \leq \hat{t}_j < t_{i+1}$ and 0 otherwise, the time-step Δt being small enough that not more than one spike can occur in one time bin), and the conditional likelihood of the spikes therefore reads

$$\log p(s|u_{\text{som}}) = \sum_{i=1}^n s_i \log [r(t_i)\Delta t] + (1 - s_i) \log [1 - r(t_i)\Delta t], \quad (10)$$

where $r(t) = g(u(t))$ with $u(t) = u_{\text{som}}(t) - u_r - (\alpha * s)(t)$.

Efficient parameter estimation

Parameter estimation (i.e. u_r, k, r_0, β, η and α) is done by maximizing the joint likelihood $p(u_{\text{som}}, s)$ as a function of the parameters; see eq. (6). A desirable feature of an optimization problem is convexity of the function to be optimized (or, equivalently, concavity of the likelihood function). Even though the problem of finding optimal parameters for the AGAPE process is not convex, the learning can be done in three alternating subspaces: (1) the GP parameters are learned (u_r, k), (2) the spike shape kernel parameters are learned, and lastly (3) the spiking parameters (r_0, β and η) are learned. In each step the remaining parameters are held fixed. The cycle is repeated until the parameters reach a region where the likelihood is locally concave in the full parameter space, after which the optimization can be run in the full parameter space until it converges. Steps (2) and (3) are convex [18] for a certain class of gain functions g , including the exponential, and therefore Newton’s method can be used in these steps as well as for the final convex optimization in the full space. Step (1) is non-convex and therefore a gradient approach is used.

Validation with synthetic data

Despite this improvement in speed and tractability, the optimization is still riddled with multiple local minima which require the use of multiple random initializations. In order to show the practicality of the learning method, synthetic data of length 100 seconds ($n = 100000$) was generated with known parameters. The learning algorithm was initialized with least-squares estimates of the covariance function parameters σ_i^2 based on the empirical autocorrelation function of u_{som} and spike-shape and adaptation kernels set to zero. The algorithm converges after a few dozen iterations (taking only three minutes on an ordinary portable computer) and recovers the correct GP, spike shape, and adaptation kernels (Fig. 2C-E). All ML estimates lie within a region of two standard deviations around the ground truth, where standard deviations are estimated from the Cramér-Rao bound, which gives a lower bound of the variance of an estimator in terms of its Fisher information [19].

The model can fit *in vivo* data

The model can also fit *in vivo* traces. This is shown in detail for the example songbird HVC dataset \mathcal{D}_1 (see Materials and Methods for a detailed description of the data sets). The model manages to capture the subthreshold and suprathreshold statistics. In particular, after learning the stationary distribution of the membrane potential u is well approximated by a Gaussian (Fig. 3B) and pronounced after-hyperpolarization is seen in the spike shape kernel (Fig. 3D). The subthreshold autocorrelation structure is well reproduced by the parametric autocorrelation function k (Fig. 3C). The adaptation kernel

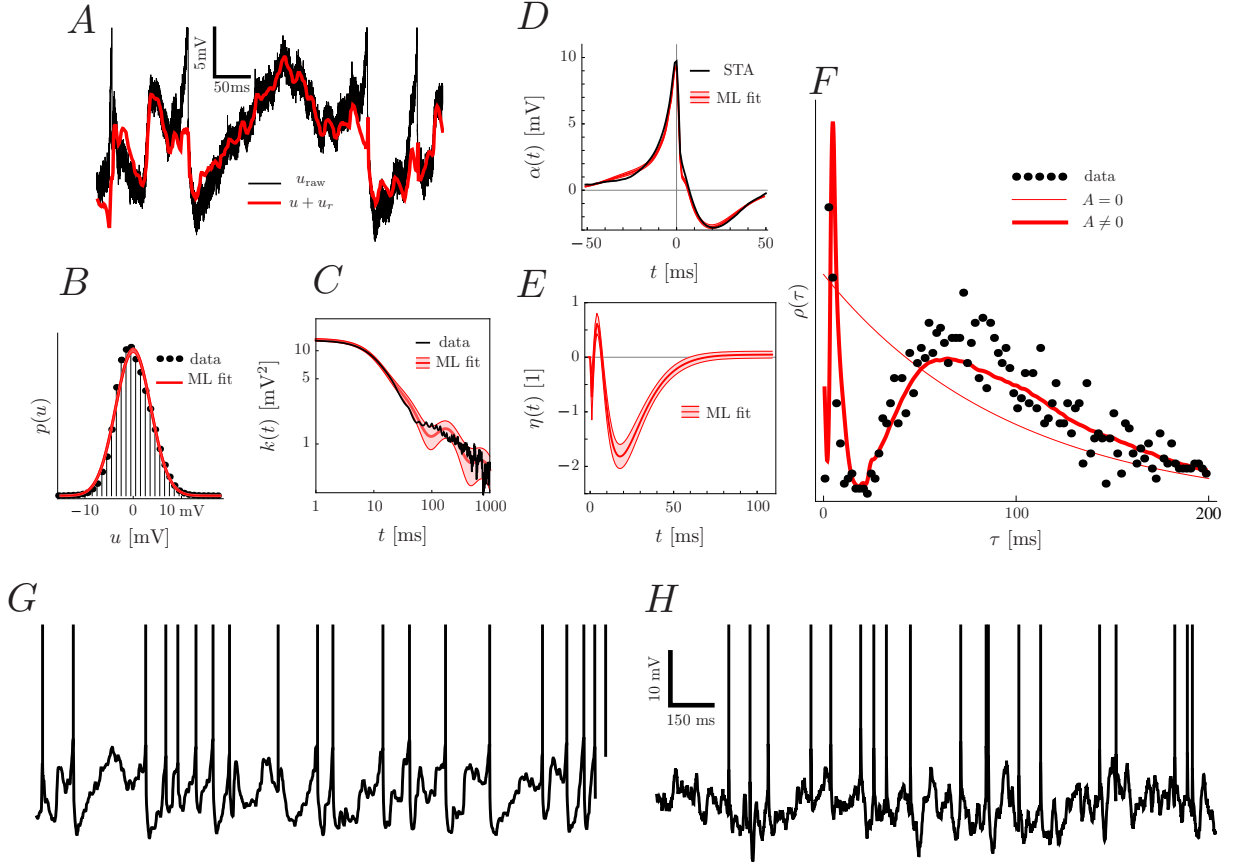


Figure 3. The results of maximum likelihood (ML) parameter fitting to dataset \mathcal{D}_1 . After learning, we see (A) the removal of the spike shape through the difference between the recorded trace u_{som} and the subthreshold membrane potential $u + u_r$; (B) the match of the stationary distribution of the subthreshold potential u and a Gaussian. We also observe that (C) the autocorrelation function of the data is well reproduced by $k(t)$ in eq. (26); (D) the spike shape kernel $\alpha(t)$ very closely matches the spike-triggered average (STA) of the membrane potential (as is expected because the coupling parameter β vanishes for this dataset); and (E) that the adaptation kernel $\eta(t)$ shows distinct modulation of firing rate which produces firing statistics significantly different from a Poisson process. This is also reflected in the inter-spike interval density $\rho(\tau)$ (F) of the data, which is well reproduced by a simulated AGAPE with adaptive kernel as in (E) (thick red line), but not by a non-adaptive (i.e. Poisson) process (thin red line). After fitting, a two second sample of synthetic data (H) looks similar as the *in vivo* data (G).

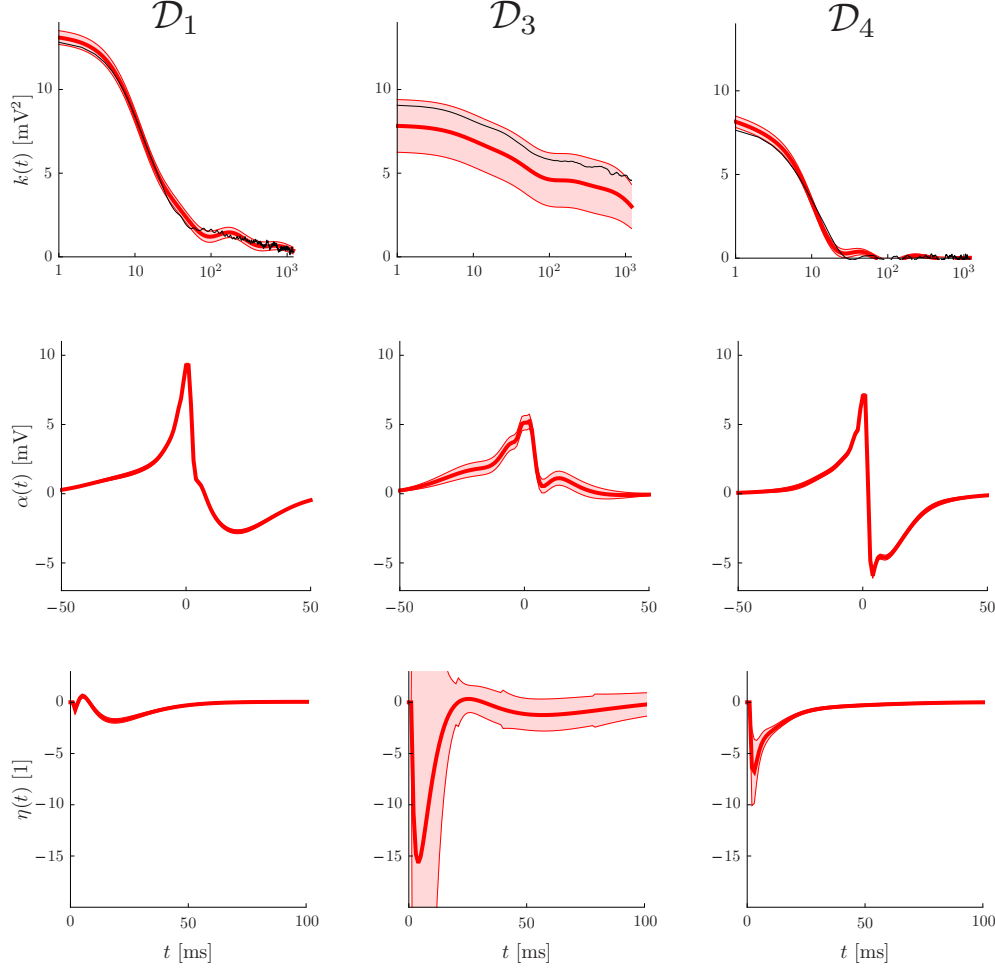


Figure 4. Fitting results for three different datasets. Dataset \mathcal{D}_1 is the same as in Fig. 3, i.e. an HVC neuron from anesthetized Zebra Finch. \mathcal{D}_3 is from HVC in awake Zebra Finch, and \mathcal{D}_4 is from mouse visual cortex in awake mouse. The different panels show the results after learning; in the first line the GP covariance function $k(t)$ (red) and the empirical autocorrelation (black), in the second line the spike shape kernel $\alpha(t)$, and in the third line the adaptation kernel $\eta(t)$. There are pronounced differences between datasets in all three kernels, showing the flexibility of the AGAPE model in describing a wide range of statistics.

reveals an interesting structure in the way the spiking statistics deviates from a Poisson process (Fig. 3E). This feature of the spiking statistics is also reflected in the inter-spike interval distribution (Fig. 3F). Both the data and the fitted model first show an increased, and then a significantly decreased probability density when compared to a pure Poisson process. The remaining parameters are $u_r = -53.3 \pm 0.2$ mV, $r_0 = 11.6 \pm 0.6$ Hz and $\beta < 1.2 \cdot 10^{-2}$ mV $^{-1}$ (errors denote two standard deviations, estimated from Fisher information, see Materials and Methods). After fitting, the parameters can be used to generate synthetic data, which is shown in Fig. 3H.

The fitting was repeated with two more datasets, \mathcal{D}_3 from another HVC neuron and \mathcal{D}_4 from mouse visual cortex, in order to show the generality of the model. The comparison of the GP, spike-shape and adaptation kernels is shown in Fig. 4. The three cells show pronounced differences in autocorrelation structure, spike shape and strength and structure of the adaptation mechanism. In particular the two datasets from songbird HVC (\mathcal{D}_1 and \mathcal{D}_3) show rather long autocorrelation lengths of the membrane potential and asymmetric spike shapes, whereas the mouse cell in \mathcal{D}_4 has comparatively short autocorrelation length and very pronounced hyperpolarization action potential shape. Regarding adaptation, again the HVC recordings show similar features (namely an inflection and positive values of the adaptation kernel), albeit on different time-scales and with different magnitude (adaptation being much stronger in \mathcal{D}_3 than in \mathcal{D}_1). Dataset \mathcal{D}_4 shows a comparatively simple form of refractoriness.

The model does not overfit *in vivo* data

The AGAPE process is fairly complex and has a fairly large number of parameters. Therefore it is important to check whether the model is too complex and overfits the data, compromising its generalization performance. In short, when a model has too many parameters, they tend to be poorly constrained by the data and therefore when the model is first trained on one part of the data and then tested on another part on which it is not trained, the test performance will be significantly worse than the training performance.

Here, we use cross-validation to perform a factorial model comparison on an exemplary dataset in order to validate the different structural parts of the model. The procedure is described in detail in the Materials and Methods.

Model comparison is performed on the dataset \mathcal{D}_2 and the results are shown in Fig. 5, where the mean difference of per-bin log-likelihood (c.f. ‘Materials and Methods’)

$$\Delta p_i^{\text{valid}} = \langle \Delta p_{ij}^{\text{valid}} \rangle_j, \quad \Delta p_i^{\text{test}} = \langle \Delta p_{ij}^{\text{test}} \rangle_j, \quad \Delta p_{ij}^{\text{valid,test}} = p_{i,j}^{\text{valid,test}} - p_{G\alpha\beta\eta,j}^{\text{valid,test}} \quad (11)$$

is shown for all $i \in \{0, \dots, G\alpha\beta\eta\}$ (here, $\langle \cdot \rangle_j$ denote averages over j). The results are very similar for both validation data (which was left-out during training, but appeared in other training runs) and the test data which was never seen during training. The most complex model ($\mathcal{M}_{G\alpha\beta\eta}$) performs significantly better than any one of the simpler models on validation data except $\mathcal{M}_{G\alpha\eta}$ where the difference is too small and lies inside a region of two standard errors of the mean. This confirms that most of the added model complexity is not superfluous, and indeed beneficial for a faithful description of the statistics.

The model captures distinct dynamical features of spontaneous and evoked activity

Several studies have reported a surprising similarity between spontaneous activity and average evoked activity [8, 20, 7]. Those results have been obtained in visual cortex of cat [7] and ferrets [8, 20]. Here we looked at recordings from primary auditory cortex of gerbils during spontaneous activity and during the presentation of an auditory stimulus (dataset \mathcal{D}_5). As can be seen in Fig. 6A, the snippets of intracellular recording look very similar between the two conditions. Here, we want to exemplify how the AGAPE

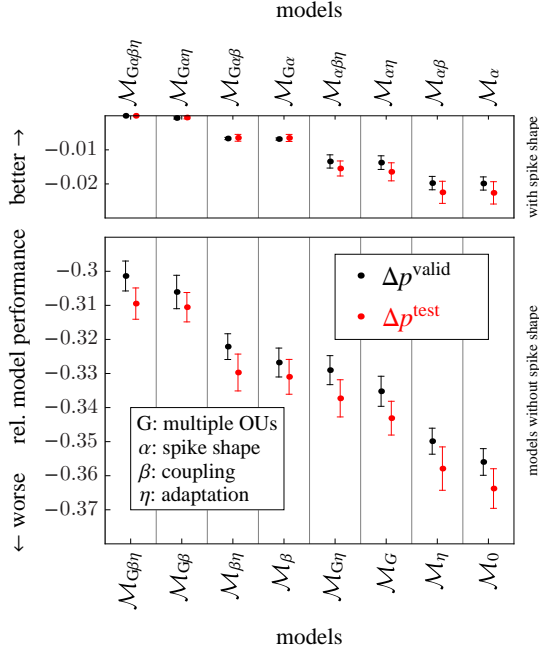


Figure 5. Comparison of the different models on dataset \mathcal{D}_2 . The relative measure of model performance, i.e. the per-bin log-likelihood Δp (c.f. eq. (11)) between any model and the most complex model ($\mathcal{M}_{G\alpha\beta\eta}$) are significantly negative (with exception of $\mathcal{M}_{G\alpha\eta}$, and trivially $\mathcal{M}_{G\alpha\beta\eta}$), implying that the added complexity improves the model fit without overfitting. This holds for both validation scores Δp^{valid} (black) and scores from unseen test data Δp^{test} (red). Error bars denote one standard error of the mean (S.E.M.). The biggest improvement of fit quality is achieved by including the spike shape kernel (upper vs. lower part of the figure).

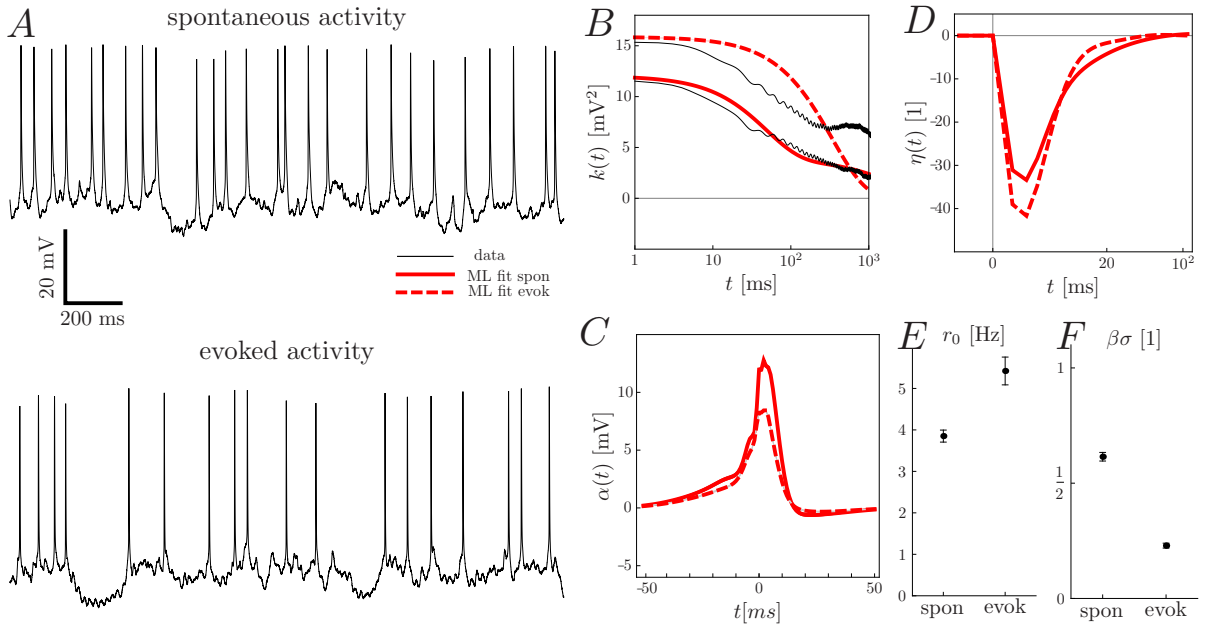


Figure 6. (A) Two second-long snippets of intracellular activity are shown for both the spontaneous and the stimulus-evoked condition. (B-D) The autocorrelation, spike shape and adaptation kernels show significant differences (i.e. they exceed the size of error margins). The stimulus-evoked activity shows a higher variance, longer autocorrelation length, smaller spike shape and more pronounced adaptation. (E-F) The baseline firing rate and the dimensionless coupling are also significantly different between the two conditions.

model can help understand the differences between the statistical properties of spontaneous and stimulus-evoked activity. We took data from auditory cortex both during spontaneous activity and during the presentation of a speech stimulus. We then fitted the model separately to the spontaneous data and the stimulus response. The spontaneous trace had a length of 238.4 seconds. For the evoked activity, there were 255 repetitions of the same 2.5 second stimulus, and we treated them as independent samples from the same distribution, yielding only one set of parameters for the whole set of responses.

The parametrization choices were the same except for the covariance kernel $k(t)$. Here, we parametrized it as a sum of two exponential functions,

$$k(t) = \sigma_1^2 e^{-\theta_1 |t|} + \sigma_2^2 e^{-\theta_2 |t|} \quad (12)$$

where θ_i are variable time-constants (in contrast to the other datasets, where we used ten exponentials with fixed time-constants, c.f. Materials and Methods).

The results of the fitting are shown in Figs. 6B-F. Firstly, the subthreshold potential variances $k(0)$ of the two conditions are different, with a value of 12 mV² for the spontaneous vs. 15.9 mV² for the evoked activity. Also the baseline firing rate r_0 is different; 3.85 Hz for the spontaneous and 5.42 Hz for the evoked activity. The coupling strength $\beta\sigma$ in units of the subthreshold standard deviation, is higher during spontaneous activity (0.62) vs. evoked (0.23).

Coming to the kernels, the differences are more subtle, but still significant. The autocorrelation lengths $1/\theta_i$ are very distinct for spontaneous activity (48 ms and 2.14 s) and very close for evoked activity (330 and 353 ms). Also note that in Fig. 6B the empirical autocorrelation (black line) does not perfectly coincide with the ML estimate of the covariance function $k(t)$. This might be due to the fact that the evoked activity recordings are made of short chunks (2.5s long) and this induces finite size effects. The spike shape kernels are very similar in shape, but the one corresponding to spontaneous activity has a significantly larger amplitude. For the adaptation kernel, the two datasets are also close, but the adaptation is slightly longer-lasting for the spontaneous activity.

Discussion

In this study, we introduced the AGAPE generative model for single-neuron statistics in order to describe the spontaneous dynamics of the somatic potential without reference to an input current. We showed that this model has a rich repertoire of dynamics and can still be learned efficiently. Finally, we showed that this model can be fitted to various intracellular *in vivo* recordings while proving that the model is not overfitting, i.e. it is not too complex for the available data.

Flexibility and tractability of the model

The AGAPE model provides a unified description of intracellular dynamics, offering a large degree of flexibility in accounting for the distinct statistical features of a neuron. As has been shown for the example datasets, the model readily teases apart the differences in the statistics which exist between different cells in different animals (c.f. Fig. 4), and even between slightly different recording conditions in the same cell (c.f. Fig. 6). This shows that the model is sensitive enough to distinguish between datasets which are in fact very similar.

We used a set of approximations and techniques to make the model fitting tractable, despite its non-convexity. It is still the case that multiple local maxima of the likelihood function can make the fitting somewhat hard, especially if the quantity of data available for fitting is quite low. However, since one run of the fitting itself takes only a few minutes even on a portable computer, multiple initializations can be tried out in a relatively short amount of time.

Comparison with existing models

From an operational perspective, existing spiking neuron models can be divided into three main categories: stimulus-driven, current-driven and input-free spiking neurons. The first category contains phenomenological models of spiking neurons that aim to relate sensory stimuli to the spiking output of the neuron. The linear-nonlinear-Poisson model (LNP) [21] or the generalized linear model (GLM) [22] are typical examples in this category. Even though the spiking generation of the AGAPE shares some similarities with those models, there is an important distinction to make. In those models the convolved input (c.f. the ‘L’ step of the LNP or the GLM) is an internal variable that does not need to be identified to the somatic membrane potential whereas in our case, the detailed modeling of the membrane potential dynamics is an important part of the AGAPE. Furthermore, those phenomenological models are interested on extracellular spiking recordings whereas the AGAPE models the dynamics of the full membrane potential accessible with intracellular methods.

The second class of spiking models aim at bridging the gap between the input current and the spiking output. The rather simple integrate-and-fire types of models such as the exponential integrate-and-fire [23] or the spike-response model [24, 25] as well as the more biophysical models such as the Hodgkin-Huxley model [26] fall in this category. In contrast with those models where the action potentials are caused by the input current, the AGAPE produces a fluctuating membrane potential and stochastic spike generation without a reference to an input current.

The last category of spiking neurons aim at producing spontaneous spiking activity without an explicit dependence to a given input. For example, Cunningham et al. propose a doubly stochastic process where the spiking generation is given by a gamma interval process and the firing intensity by a rectified Gaussian process, which provides a flexible description of the firing statistics [27]. However, the membrane potential dynamics is not modeled. In opposition, the neuronal dynamics assumed by Pfister et al. models explicitly the membrane potential (as a simple Ornstein-Uhlenbeck process) but is not flexible enough to capture the dynamics of *in vivo* recordings [9]. Also any of the current-driven spiking neurons mentioned above can be turned into an input-independent model by assuming some additional input noise. So why is there a need to go beyond stochastic versions of those models? An integrate-and-fire model with additive Gaussian white noise is certainly learnable but does not have the flexibility to model arbitrary autocorrelation for the membrane potential. At the other end of the spectrum, a Hodgkin-Huxley model with some colored noise would certainly be able to model a richer dynamical repertoire, but the fitting of it remains challenging [10] (but see [11]). The main advantage of the AGAPE is that it is at the same time very flexible and easily learnable. The flexibility mostly comes from the fact that any covariance function can be assumed for the GP process. The learnability comes from the cumulant approximation as well as from the presence of convex subspaces in the parameter space.

Another distinct feature of our model with respect to other existing models is the explicit modeling of the spike shape. Why is this element required in the model? Like other formal spiking neuron models, the AGAPE separates the neuronal activity into two distinct features: the subthreshold membrane potential and the action potential. Since the model critically depends on the (subthreshold) membrane potential at the time of the spike (which is a quantity used in the model but that cannot be measured explicitly), it is important to describe precisely how those two quantities are separated. A first possibility would be to compute the spike-triggered average of the recorded voltage trace and remove it from the raw data. However, this procedure also removes the upward fluctuation of the subthreshold potential which caused the spike. In contrast, the parametric approach which is used here allows the model to find the optimal balance between removing the stereotypical spike waveform while retaining the subthreshold fluctuation which is consistent with the GP and predictive of the spike. For data where the spike shape shows considerable variability, the model could be generalized to include a stochastic spike shape kernel or a history dependent spike shape kernel.

Extensions and future directions

Despite the focus on the present work on single-neuron spontaneous dynamics, the AGAPE model admits a straightforward inclusion of both stimulus-driven input and recurrent input. The inclusion of stimulus-driven input is similar as for the GLM model and allows the model to capture the neuronal correlate of stimulus-specific computation. The recurrent input makes the framework adaptable to multi-neuron recordings *in vivo*. While intracellular recordings from many neurons *in vivo* are very hard to perform, the rapid development of new recording techniques (e.g. voltage-sensitive dyes) makes the future availability of subthreshold data with sufficient time-resolution at least conceivable. The full-fledged model would allow questions regarding the relative importance of background activity, recurrent activity due to computation in the circuit, and activity directly evoked by sensory stimuli to be answered in a systematic way. In this setup, the contribution of the GP distributed membrane potential to the overall fluctuations would be reduced (since it has to capture less unrecorded neurons) while the contribution of the recorded neurons would increase. This modified model can be seen as a generalization of the stochastic spike-response model [24] or a generalization of the GLM (if the internal variable of the GLM is interpreted as the membrane potential).

In the present study, the AGAPE was fit to different datasets of two different animals and brain regions. A systematic fitting to *in vivo* intracellular data from a wide range of animals and brain regions would constitute a classification scheme which does not only complement existing classifications of neurons which are based on electrophysiological, morphological, histological, and biochemical data; such as the one in [28], but which is in direct relationship with the computational tasks the brain is facing *in vivo*.

Another application of the AGAPE could be in the context of a normative theory of short-term plasticity. Indeed, it has been recently proposed that short-term performs Bayesian inference of the presynaptic membrane potential based on the observed spike-timing [9, 29]. According to this theory, short-term plasticity properties have to match the *in vivo* statistics of the presynaptic neuron. The AGAPE provides a realistic generative model of presynaptic statistics under which to do the inference and thereby derive predictions on the short-term plasticity properties of downstream synapses.

Materials and Methods

Description of the datasets used

1. Dataset \mathcal{D}_1 is a recording from a HVC neuron of an anesthetized Zebra Finch (Ondracek and Hahnloser, unpublished recordings). The recording has a total length of 270 seconds at 32 kHz (c.f. Fig. 1A for a snippet of this recording) and contains 2281 action potentials.
2. Dataset \mathcal{D}_2 is another recording from a projection cell in HVC of Zebra Finch, but this time the animal is awake (Vallentin and Long, unpublished recordings). It consists of 6 individual recordings which together have a length of 152.5 seconds at 40 kHz. This dataset is used for model comparison (see below).
3. Dataset \mathcal{D}_3 is from similar conditions as \mathcal{D}_2 (Vallentin and Long, unpublished recordings) and has a length of 60 seconds.
4. Dataset \mathcal{D}_4 consists of 19 individual trials of 4.95s duration at 20 kHz. The recording was obtained from a pyramidal neuron in layer 2/3 of awake mouse visual cortex [30].
5. Dataset \mathcal{D}_5 is from primary auditory cortex (A1) of gerbils under ketamine/xylazine anaesthesia (Garcia-Lazaro and McAlpine, unpublished recordings). It consists of a 238.4 seconds-long recording of spontaneous activity, and 255 responses to repetitions of the same 2.5 s auditory stimulus. The auditory stimulus consisted of a female voice reciting: ‘A huge tapestry hung in her hallway’.

Preprocessing

Intracellular voltage traces are often recorded at a rate between 20 and 40 kHz. This allows the action potentials to be resolved very clearly and precise spike timings to be extracted. However, for the study of the subthreshold regime, this high sampling rate is not required, and therefore the data may be down-sampled to roughly one kHz after obtaining the precise spike timings. This reduces the computational demands during data fitting.

We define the spike times \hat{t}_i operationally as the time where the local maximum of the action potential is reached. This means that \hat{t}_i occurs after action potential onset, and hence the spike shape kernel has to extend to the past of \hat{t}_i . Other definitions of \hat{t}_i are of course possible, but here we chose this one for its ease of determining the local maximum. The spike times \hat{t}_i are then binned to 1 ms, yielding a binary spike train $s_i = 0, 1$.

For $u_{\text{som}}(t)$ we use a preprocessed version of the recorded trace which has been median-filtered with a width of the filter of 3.75 ms and then down-sampled to 1 kHz, making it the same length as the binary spike train. This procedure preserves the relevant correlation structure of the membrane potential while reducing the computational demands of fitting as much as possible.

Circulant matrices

In order to reduce the computational complexity of the likelihood estimation, we approximate the auto-covariance matrix K (which is a Toeplitz matrix K) with a circulant matrix C . By definition a circulant matrix can be expressed as

$$C_{ij} = c_{(i-j \bmod n)+1} \quad (13)$$

we write $C = C_n(c)$. The task is now to find a circulant matrix which is as close as possible to the covariance matrix K .

Discrete Fourier transform

Before deriving an approximation of K , let us review some properties of the discrete Fourier transform. All circulant matrices of dimension n can be diagonalized by the unitary discrete Fourier transform matrix \mathbb{U}

$$C_n(c) = \mathbb{U}_n^\dagger \text{diag}(\mathbb{F}_n c) \mathbb{U}_n, \quad \mathbb{U}_n = \frac{1}{\sqrt{n}} \mathbb{F}_n, \quad (\mathbb{F}_n)_{ij} = e^{\frac{2\pi I(i-1)(j-1)}{n}} \quad (14)$$

where I denotes the imaginary unit, \dagger is the conjugate transpose and the diagonal matrix contains the discrete Fourier transform of the defining vector c , $\hat{c} = \mathbb{F}_n c$. The latter is the vector of eigenvalues of C , and it is real. This fact, combined with the power of Fast Fourier Transforms (FFT) makes calculation of inverse and determinant of C extremely cheap, as is multiplication of C^{-1} by a vector $x \in \mathbb{R}^n$, which simplifies to

$$C_n^{-1}(c)x = \mathbb{U}_n^\dagger \left(\frac{\mathbb{U}_n x}{\mathbb{F}_n c} \right) \quad (15)$$

where the vector in brackets is the component-wise quotient of the vectors $\mathbb{U}_n x$ and $\mathbb{F}_n c$. Inner products of $x, y \in \mathbb{R}^n$ defined by the inverse covariance turn into

$$y^T C_n^{-1}(c)x = (\mathbb{U}_n y)^\dagger \left(\frac{\mathbb{U}_n x}{\mathbb{F}_n c} \right) \quad (16)$$

This result is the key to express in a computationally convenient way the log probability of the membrane potential u_{som} , $\log p(u_{\text{som}})$ (see eq. (9)). With these tools, we can now tackle the following minimization problem

Circulant approximation

The task is now to find a circulant matrix which is as close as possible to the covariance matrix K . This can be formalized as the following minimization problem:

$$C = \underset{D \text{ circulant}}{\operatorname{argmin}} D_{\text{KL}}(\mathcal{N}(m, K) || \mathcal{N}(m, D)), \quad \forall m \quad (17)$$

where \mathcal{N} denotes a multivariate Gaussian with specified mean vector and covariance matrix. This problem has the unique solution

$$c_i = \frac{1}{n} \{(n - i + 1)k_i + (i - 1)k_{n-i+2}\}, \quad 1 \leq i \leq n, \quad k_{n+1} \equiv 0 \quad (18)$$

Proof: The Kullback-Leibler divergence between two Gaussians is given by

$$D_{\text{KL}}(\mathcal{N}(m, K) || \mathcal{N}(m, C)) = \frac{1}{2} [\operatorname{tr}(C^{-1}K) - \log \det(C^{-1}K)] - \frac{n}{2} \quad (19)$$

We have

$$C^{-1} = \mathbb{U}_n^\dagger \operatorname{diag} \left(\frac{1}{\hat{c}} \right) \mathbb{U}_n, \quad \det C^{-1} = \prod_{i=1}^n \frac{1}{\hat{c}_i} \quad (20)$$

and hence

$$D_{\text{KL}}(\mathcal{N}(m, K) || \mathcal{N}(m, C)) = \frac{1}{2} \sum_{i=1}^n \left[\frac{(\mathbb{U}_n K \mathbb{U}_n^\dagger)_{ii}}{\hat{c}_i} + \log \hat{c}_i \right] + \text{const.} \quad (21)$$

where the constant does not depend on c . We obtain the derivative

$$\frac{\partial}{\partial c_i} D_{\text{KL}}(\mathcal{N}(m, K) || \mathcal{N}(m, C)) = \frac{1}{2\hat{c}_i} \left[1 - \frac{(\mathbb{U}_n K \mathbb{U}_n^\dagger)_{ii}}{\hat{c}_i} \right] = 0 \quad (22)$$

and therefore, at the stationary point we have

$$\begin{aligned} \hat{c}_i &= (\mathbb{U}_n K \mathbb{U}_n^\dagger)_{ii} \\ c_i &= \frac{1}{n^2} \sum_{j,l,m=1}^n (\mathbb{F}_n^\dagger)_{ij} (\mathbb{F}_n)_{jl} K_{lm} (\mathbb{F}_n^\dagger)_{mj} \\ &= \frac{1}{n^2} \sum_{j,l,m=1}^n K_{lm} \exp \left[\frac{2\pi I}{n} (-(i-1)(j-1) + (j-1)(l-1) - (m-1)(j-1)) \right] \\ &= \frac{1}{n^2} \sum_{j,l,m=1}^n k_{|l-m|+1} \exp \left[\frac{2\pi I}{n} (j-1)(l-m+1-i) \right] \end{aligned} \quad (23)$$

The sum of roots of unity over j only gives a non-zero value if the integer $q = l - m + 1 - i$ is a multiple of n . Since q has a maximum of $q = n - 1$ when $l = n, m = i = 1$ and a minimum of $q = 2 - 2n$ when $l = 1, m = i = n$, only $q = 0$ and $q = -n$ are eligible. Hence,

$$\begin{aligned} c_i &= \frac{1}{n^2} \sum_{l,m=1}^n n k_{|l-m|+1} (\delta_{0,l-m+1-i} + \delta_{-n,l-m+1-i}) \\ &= \frac{1}{n} \sum_{r=-n+1}^{n-1} (n - |r|) k_{|r|+1} (\delta_{i-1,r} + \delta_{i-1,n+r}) \\ &= \frac{1}{n} \{(n - i + 1)k_i + (i - 1)k_{n-i+2}\}, \quad k_{n+1} \equiv 0 \end{aligned} \quad (24)$$

The second equality is obtained by reparametrizing $r = l - m$. \square

Parametrizations and initializations

We already introduced the parameters u_r , r_0 and β . Additional parameters are needed to describe the autocorrelation $k(t)$, the spike shape $\alpha(t)$ and the adaptation kernel $\eta(t)$.

The covariance function of the GP has to be parametrized such that it can explain the autocorrelation structure of the data. Therefore, an initial examination of the empirical autocovariance of u_{som} , i.e.

$$k^{\text{emp}}(j\Delta t) = \frac{1}{n-j-1} \sum_{i=1}^{n-j} \left(u_{\text{som},i} - \frac{1}{n-j} \sum_{k=1}^{n-j} u_{\text{som},k} \right) \left(u_{\text{som},i+j} - \frac{1}{n-j} \sum_{k=1}^{n-j} u_{\text{som},k+j} \right), \quad (25)$$

for $j = 0, \dots, j_{\text{max}}$, is done in order to determine a suitable basis. Here, we used a sum of Ornstein-Uhlenbeck (OU) kernels, i.e.

$$k(t) = \sum_{i=1}^{n_k} \sigma_i^2 e^{-\theta_i |t|}, \quad (26)$$

where $n_k = 10$ and $\theta_i = 2^{-i} \text{ ms}^{-1}$. The autocovariance has to remain positive definite. This induces the following linear constraints:

$$\hat{c}_i = \sum_{j=1}^{n_k} \sigma_j^2 \hat{c}_i^{(j)} > 0, \quad \forall i = 1, \dots, n, \quad (27)$$

on σ_i^2 , where $\hat{c}_i^{(j)}$ are the discrete Fourier transforms of the circulant basis vectors. The optimization problem is non-convex in the subspace of σ_i^2 and multiple local maxima and saddle points can occur. Therefore, multiple initializations have to be made in order to find a potential global optimum. In general, the least-squares fit of $k(t)$ to the empirical autocovariance function (25) yields a good starting point for the optimization.

The spike-rate adaptation kernel is chosen to be a linear combination of ten different alpha shapes

$$\eta(t) = \begin{cases} \sum_{i=1}^{n_\eta} w_i [\exp(-\nu_i t) - \exp(-\omega_i t)], & t > 0, \\ 0, & t \leq 0, \end{cases} \quad (28)$$

where we chose $n_\eta = 10$, $\nu_i = 2\omega_i$ and $\nu_i = 2^{-i} \text{ ms}$. Lastly, the spike shape kernel is a sum of five Gaussians and five derivatives of Gaussians, i.e.

$$\alpha(t) = \sum_{i=1}^{10} a_i \alpha^{(i)}(t), \quad \alpha^{(i)}(t) = e^{-\frac{1}{2}t^2/\tau_i^2}, \quad \alpha^{(i+5)}(t) = t e^{-\frac{1}{2}t^2/\tau_i^2}, \quad i = 1, \dots, 5, \quad (29)$$

where the time-constants τ_i are 1,2,5,10,20 ms. The spike-shape and adaptation kernels as well as u_r , r_0 and β have to satisfy the constraints

$$0 < r(t_i)\Delta t = r_0 \Delta t e^{\beta[u_{\text{som},i} - u_r - (\alpha * s)_i] + (\eta * s)_i} < 1, \quad \forall i = 1, \dots, n. \quad (30)$$

The constraint $\beta \geq 0$ is also used in order to have the property that the firing rate is higher when the cell is more depolarized.

Model validation

We performed a factorial model comparison (see Fig. 5) where the four factors were the presence/absence of each of the following: multiple OU components in the GP autocorrelation function (c.f. eq. (26), as opposed to only one OU kernel with variable time-constant), the spike shape α , coupling between u and s (through β) and adaptation η , which gives a total of 16 different models. We use the nomenclature that

\mathcal{M}_0 is the simplest model, e.g. $\alpha = \beta = \eta = 0$ and only one OU component, having only four parameters (u_r, θ, σ and r_0). A subscript G (for GP) indicates that we use the multiple OU basis and any other subscript indicates that the corresponding parameter is adjustable in addition to the parameters already present in \mathcal{M}_0 and the parameters that are associated with the subscribed ones. E.g. $\mathcal{M}_{G\alpha}$ indicates that we use the multiple OU basis and allow a non-zero spike-shape kernel and that there are now 22 parameters (u_r, θ_i, a_i for $i = 1, \dots, 10$ and r_0).

For each of the models $\mathcal{M} \in \{\mathcal{M}_0, \dots, \mathcal{M}_{G\alpha\beta\eta}\}$, we performed eight-fold leave-one-out cross-validation [31] on dataset \mathcal{D}_2 in order to assess the models' generalization performance. The entire dataset was cut into eight equally-sized chunks d_j , where $j = 1, \dots, 8$, each of length 15s ($n = 15000$), and six chunks of 3s d'_j , $j = 1, \dots, 6$ set aside for validation ($n' = 3000$). Each model was then trained on seven out of eight chunks (treating them as independent samples) giving an optimal set of parameters $\Theta_j^i = \arg\max_{\Theta} p(\{d_k, k \neq j\} | \mathcal{M}_i, \Theta)$ and training per-bin log-likelihood $p_{ij}^{\text{train}} = \frac{1}{7n} \log p(\{d_k, k \neq j\} | \mathcal{M}_i, \Theta_j^i)$. Then the validation likelihood $p_{ij}^{\text{valid}} = \frac{1}{n} \log p(d_j | \mathcal{M}_i, \Theta_j^i)$ of the left-out chunk $\#j$ was evaluated. The unseen data d'_j is used for a final benchmark of model performance, where the best set of parameters is selected for each model, i.e. $p_{ij}^{\text{test}} = \frac{1}{n'} \max_{k=1, \dots, 8} \log p(d'_j | \mathcal{M}_i, \Theta_k^i)$.

Cramér-Rao Bounds

The variance of a maximum likelihood estimator can be difficult to estimate. The Cramér-Rao bound serves to give a lower bound to the variance of such an estimator in terms of the Fisher information matrix. Denote by $\Theta = (\theta_1, \dots, \theta_q)$ the vector of all parameters, which live in a parameter space \mathcal{S} and let $\Theta \mapsto p(X|\Theta)$ be the likelihood function of the model for state X . The Fisher information matrix is defined as

$$\mathcal{J}_{ik}(\Theta) = \int dX p(X|\Theta) \frac{\partial \log p(X|\Theta)}{\partial \theta_i} \frac{\partial \log p(X|\Theta)}{\partial \theta_k} = - \int dX p(X|\Theta) \frac{\partial^2 \log p(X|\Theta)}{\partial \theta_i \partial \theta_k}, \quad (31)$$

where the integral is over the whole state space of X . Let $f : \mathcal{S} \rightarrow \mathbb{R}$ be a real-valued function of the parameters and $T(X)$ be an estimator of that function, e.g. using the maximum likelihood estimator of the model parameters:

$$T(X) = f \left(\arg\max_{\Theta \in \mathcal{S}} p(X|\Theta) \right). \quad (32)$$

Moreover, let $\psi(\Theta) = E[T(X)]$ be the expectation of the estimator. Note that in general, the estimator is biased, i.e. $b(\Theta) = \psi(\Theta) - f(\Theta) \neq 0$. Under some regularity conditions [32], the variance of the estimator T satisfies

$$\text{Var}[T(X)] \geq \psi'(\Theta)^\top \mathcal{J}^{-1}(\Theta) \psi'(\Theta), \quad (33)$$

where $\psi'(\Theta)$ is the gradient of ψ . This is called the Cramér-Rao bound. Because of

$$E[(T(X) - f(\Theta))^2] = \text{Var}[T(X)] + E[b^2(\Theta)] \geq \text{Var}[T(X)], \quad (34)$$

the bound holds also for the mean-squared error (MSE) of the function f .

To exemplify, let's assume we want to estimate the error of the covariance function $k(t)$ of the GP. We have a family of functions $f_t(\Theta)$ and corresponding estimators

$$T_t(X) = f_t(\Theta_{\text{ML}}(X)), \quad \Theta_{\text{ML}}(X) = \arg\max_{\Theta \in \mathcal{S}} p(X|\Theta). \quad (35)$$

Since $\psi_t(\Theta)$ is not known, we have to use (33) with ψ replaced by f . Provided that the bias $b(\Theta)$ varies slowly compared to ψ itself, this will yield a good approximation. Therefore, upon observing the data D , we set $\hat{\Theta} = \Theta_{\text{ML}}(D)$ and $k(t) = T_t(\hat{\Theta})$. We now want to estimate by how much T_t is scattered around $k(t)$ if we repeatedly apply it to synthetic data with parameters $\hat{\Theta}$. Equation (33) can be used with ψ_t replaced by $k(t)$ and yields a lower bound to the variance of our estimate.

Acknowledgments

We would like to thank Janie Ondracek and Richard Hahnloser in Zürich, Switzerland, Daniela Vallentin and Michael Long at NYU, Bilal Haider and Matteo Carandini, Jose Garcia-Lazaro and David McAlpine at UCL for kindly providing the data for this study.

References

1. Woody CD, Gruen E (1978) Characterization of electrophysiological properties of intracellularly recorded neurons in the neocortex of awake cats: a comparison of the response to injected current in spike overshoot and undershoot neurons. *Brain research* 158: 343–357.
2. Baranyi A, Szente MB, Woody CD (1993) Electrophysiological characterization of different types of neurons recorded in vivo in the motor cortex of the cat. II. Membrane parameters, action potentials, current-induced voltage responses and electrotonic structures. *Journal of neurophysiology* 69: 1865–1879.
3. Steriade M, Timofeev I, Grenier F (2001) Natural waking and sleep states: a view from inside neocortical neurons. *Journal of neurophysiology* 85: 1969–1985.
4. Matsumura M, Cope T, Fetz EE (1988) Sustained excitatory synaptic input to motor cortex neurons in awake animals revealed by intracellular recording of membrane potentials. *Experimental Brain Research* 70: 463–469.
5. Poulet JFA, Petersen CCH (2008) Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature* 454: 881–885.
6. Lee AK, Manns ID, Sakmann B, Brecht M (2006) Whole-Cell Recordings in Freely Moving Rats. *Neuron* 51: 399–407.
7. El Boustani S, Marre O, Béhuret S, Baudot P, Yger P, et al. (2009) Network-State Modulation of Power-Law Frequency-Scaling in Visual Cortical Neurons. *PLoS Computational Biology* 5: e1000519.
8. Berkes P, Orban G, Lengyel M, Fiser J (2011) Spontaneous Cortical Activity Reveals Hallmarks of an Optimal Internal Model of the Environment. *Science* 331: 83–87.
9. Pfister JP, Dayan P, Lengyel M (2010) Synapses with short-term plasticity are optimal estimators of presynaptic membrane potentials. *Nature Neuroscience* 13: 1271–1275.
10. Gerstner W, Naud R (2009) How Good Are Neuron Models? *Science* 326: 379–380.
11. Druckmann S, Banitt Y, Gidon A, Schürmann F, Markram H, et al. (2007) A novel multiple objective optimization framework for constraining conductance-based neuron models by experimental data. *Frontiers in neuroscience* 1: 7–18.
12. Møller J, Syversveen A, Waagepetersen R (1998) Log gaussian cox processes. *Scandinavian Journal of Statistics* 25: 451–482.
13. Rasmussen CE, Williams CK (2006) *Gaussian Processes for Machine Learning*. the MIT Press.
14. Buzsáki G (2002) Theta oscillations in the hippocampus. *Neuron* 33: 325–340.
15. Shinomoto S, Tsubo Y (2001) Modeling spiking behavior of neurons with time-dependent Poisson processes. *Physical Review E* 64: 041910.

16. Softky WR, Koch C (1993) The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *The Journal of Neuroscience* 13: 334–350.
17. Gray RM (2006) Toeplitz and circulant matrices: A review. *Foundations and Trends in Communications and Information Theory* 2: 155–239.
18. Paninski L (2004) Maximum likelihood estimation of cascade point-process neural encoding models. *Network: Computation in Neural Systems* 15: 243–262.
19. Rao CR (1945) Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society* 37: 81–91.
20. Fiser J, Chiu C, Weliky M (2004) Small modulation of ongoing cortical dynamics by sensory input during natural vision. *Nature* 431: 573–578.
21. Chichilnisky EJ (2001) A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems* 12: 199–213.
22. Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, et al. (2008) Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454: 995–999.
23. Brette R, Gerstner W (2005) Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of neurophysiology* 94: 3637–3642.
24. Gerstner W, Kistler WM (2002) *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. Cambridge University Press.
25. Jolivet R, Rauch A, Lüscher HR, Gerstner W (2006) Predicting spike timing of neocortical pyramidal neurons by simple threshold models. *Journal of Computational Neuroscience* 21: 35–49.
26. Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology* 117: 500.
27. Cunningham JP, Byron MY, Shenoy KV, Sahani M (2007) Inferring Neural Firing Rates from Spike Trains Using Gaussian Processes. In: *Advances in neural information processing systems*. pp. 329–336.
28. Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, et al. (2004) Interneurons of the neocortical inhibitory system. *Nature Reviews Neuroscience* 5: 793–807.
29. Pfister JP, Dayan P, Lengyel M (2009) Know thy neighbour: A normative theory of synaptic depression. *Advances in neural information processing systems* 22: 1464–1472.
30. Haider B, Häusser M, Carandini M (2013) Inhibition dominates sensory responses in the awake cortex. *Nature* 493: 97–100.
31. Arlot S, Celisse A (2010) A survey of cross-validation procedures for model selection. *Statistics surveys* 4: 40–79.
32. Cramér H (1946) *Mathematical Methods of Statistics*. Princeton University Press.